

Basis Set Approach to Solution of Poisson Equation for Small Molecules Immersed in Solvent

LAURENT DAVID and MARTIN J. FIELD*

Laboratoire de Dynamique Moléculaire, Institut de Biologie Structurale—Jean-Pierre Ebel, 41 Avenue des Martyres, 38027 Grenoble Cedex 1, France

Received 9 November 1995; accepted 14 May 1996

ABSTRACT

The problem of how to calculate the electrostatic interactions between molecules and a solvent is a very important one in theoretical chemistry and biophysics. One of the more commonly used methods has been to represent the solvent by a dielectric continuum and then to solve the Poisson (or the Poisson–Boltzmann) equation for the potential due to the charge distribution of the solute. The solution of the equation has, up to now, been largely carried out using finite-difference grid-based methods. In this article, we investigate the use of an alternative method, based on a basis set expansion of the potential. The choice of basis functions, the representation of the dielectric function and the accuracy that can be obtained are discussed and illustrated by example calculations on small molecules. © 1997 by John Wiley & Sons, Inc.

Introduction

The behavior and properties of molecules are profoundly changed when they are immersed in solution. These changes are due, in large part, to the electrostatic interactions between the solute and the solvent, and so an ability to calculate these terms is essential if the effects of solvent are to be understood. One of the more successful models

which has been used to study solute/solvent interactions (particularly for large macromolecular systems, such as proteins) has been to describe the solvent as a dielectric continuum and to solve the Poisson–Boltzmann equation for the electrostatic potential, and hence the energy of solvation, due to the charge distribution in the solute.^{1–4}

Up to now, the methods of choice for the solution of the Poisson–Boltzmann equation have been the numerical finite difference⁴ and, to a lesser extent, finite element⁵ and boundary element approaches.^{6,7} These have been developed to the point where they can be used to rapidly obtain the

*Author to whom all correspondence should be addressed.
E-mail: mjfield@ibs.fr

electrostatic potential for systems containing several thousands of atoms. The state-of-the-art is such that the Poisson–Boltzmann equation is starting to be used to calculate the electrostatic energy of a system and the forces on its constituent particles during molecular dynamics simulations (see, e.g., ref. 8). However, these algorithms have some limitations. The precision of the finite difference methods, for example, is limited by the size of the grid that can be employed. For large systems the use of a fine grid is expensive and so a coarser grid must be used with a consequent reduction in the accuracy of the potential that can be obtained.

In this article we explore an alternative method to the solution of the Poisson equation for a solute/solvent system. It is based on the expansion of the electrostatic potential in terms of a linear combination of basis functions. While basis set methods have been used widely for the solution of other partial differential equations of interest in chemical physics, such as the Schrödinger equation,⁹ they have not been used, to our knowledge, for this application. There are several possible advantages to using a basis set. These include the fact that the potential, when expressed in terms of a basis set, has an analytic form and its derivatives and those of the solvation energy are straightforward to evaluate—a fact which is important if continuum models are to be useful for molecular dynamics simulations. On the other hand, one of the disadvantages of basis set methods is that many basis functions may be needed to provide a reasonable approximation of the potential in the system.

The outline of this article is as follows: the next section describes the basis set expansion approach to the solution of the Poisson equation; the third section presents the results of calculations on some model systems; and the last section concludes.

Methods

INTRODUCTION OF A BASIS SET

In this article we consider only the Poisson equation. The addition of the term dependent on the ionic strength is relatively straightforward (at least in its linear form) and will be considered in future work.

The Poisson equation determines the electrostatic potent, $\phi(\mathbf{r})$, as a function of the position coordinate, \mathbf{r} , in a system:

$$\nabla \cdot [\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] + 4\pi\rho(\mathbf{r}) = 0 \quad (1)$$

where ρ is the charge distribution and ε is the dielectric constant. Both of these quantities, like the potential, also depend upon the position coordinate.

It has been shown that there exists an expression which represents a free energy functional, G , for the Poisson equation. It is^{10–13}

$$G = \int d\mathbf{r} \left(\rho(\mathbf{r})\phi(\mathbf{r}) - \frac{\varepsilon(\mathbf{r})}{8\pi} (\nabla\phi(\mathbf{r}))^2 \right) \quad (2)$$

The Poisson equation results from the functional variation of this expression for the electrostatic free energy with respect to the potential. In other words, the electrostatic free energy is at an extremum for the potential which obeys the Poisson equation.

Eq. (2) defines the *variational principle* that we can use to obtain the electrostatic potential using a basis set approach. To do this, we assume that the potential can be expanded as a linear combination of M potential basis functions, $\{f_i\}$:

$$\phi(\mathbf{r}) = \sum_{i=1}^M c_i f_i(\mathbf{r}) \quad (3)$$

where the $\{c_i\}$ are the expansion coefficients for the potential. Substituting the expression for the potential into the expression for the electrostatic energy [eq. (2)] gives the matrix equation:

$$G = \sum_i c_i P_i - \sum_i \sum_j c_i c_j F_{ij} \quad (4)$$

where P_i and F_{ij} are integrals over the basis functions which are defined as follows:

$$P_i = \int d\mathbf{r} \rho(\mathbf{r}) f_i(\mathbf{r}) \quad (5)$$

$$F_{ij} = \int d\mathbf{r} \frac{\varepsilon(\mathbf{r})}{8\pi} (\nabla f_i(\mathbf{r})) \cdot (\nabla f_j(\mathbf{r})) \quad (6)$$

Before we can minimize this expression with respect to the expansion coefficients we need to take account of the fact that the total charge on the system, q , is fixed. That is:

$$\int d\mathbf{r} \rho(\mathbf{r}) = q \quad (7)$$

This equation constrains the value of the expansion coefficients because the charge density determines the value of the potential through the Poisson equation. Substituting the expression for the charge density from eq. (1) and for the potential

from eq. (3) into eq. (7) gives a constraint condition for the expansion coefficients:

$$\sum_i c_i K_i + q = 0 \quad (8)$$

where the K_i are integrals defined as:

$$K_i = \frac{1}{4\pi} \int d\mathbf{r} \nabla \cdot [\varepsilon(\mathbf{r}) \nabla f_i(\mathbf{r})] \quad (9)$$

We now minimize the expression for the free energy [eq. (4)] with respect to the expansion coefficients, subject to the constraint condition [eq. (8)]. Solving for the vector of expansion coefficients, \mathbf{c} , gives:

$$\mathbf{c} = \frac{1}{2} \mathbf{F}^{-1} (\mathbf{P} - \lambda \mathbf{K}) \quad (10)$$

where \mathbf{F} is a matrix of the F_{ij} integrals and \mathbf{P} and \mathbf{K} are vectors of the P_i and K_i integrals, respectively. λ is a Lagrange multiplier that has been introduced to satisfy the constraint condition. It has the form:

$$\lambda = \frac{2q + \mathbf{K}^T \mathbf{F}^{-1} \mathbf{P}}{\mathbf{K}^T \mathbf{F}^{-1} \mathbf{K}} \quad (11)$$

Once the coefficients have been determined by solving these equations the energy can be calculated by substituting for the vector \mathbf{c} in eq. (4).

CALCULATION OF DERIVATIVES

In molecular simulations it is often important to be able to calculate the forces on the atoms (and hence the first derivatives with respect to the atomic positions). In the basis set approach the derivatives of the energy with respect to a parameter, x , such as the coordinate of an atom, are straightforward to obtain. Differentiating the expression for the energy gives:

$$\frac{\partial G}{\partial x} = \mathbf{c}^T \left(\frac{\partial \mathbf{P}}{\partial x} - \frac{\partial \mathbf{F}}{\partial x} \mathbf{c} \right) + (\mathbf{P} - 2\mathbf{F}\mathbf{c})^T \frac{\partial \mathbf{c}}{\partial x} \quad (12)$$

The last term can be simplified by rearranging the expression for \mathbf{c} [eq. (10)]:

$$(\mathbf{P} - 2\mathbf{F}\mathbf{c})^T \frac{\partial \mathbf{c}}{\partial x} = \lambda \mathbf{K}^T \frac{\partial \mathbf{c}}{\partial x} \quad (13)$$

It is not necessary to calculate the derivatives of the expansion coefficients with respect to x . To see this we differentiate the constraint condition [eq. (8)]:

$$\mathbf{K}^T \frac{\partial \mathbf{c}}{\partial x} + \mathbf{c}^T \frac{\partial \mathbf{K}}{\partial x} = 0 \quad (14)$$

which equates the term involving the derivatives of the expansion coefficients to one involving the derivatives of the K_i integrals. The final expression for the derivatives is:

$$\frac{\partial G}{\partial x} = \mathbf{c}^T \left(\frac{\partial \mathbf{P}}{\partial x} - \frac{\partial \mathbf{F}}{\partial x} \mathbf{c} - \lambda \frac{\partial \mathbf{K}}{\partial x} \right) \quad (15)$$

Given suitable differentiable forms for the potential basis functions, such as the ones discussed in the following subsection, the derivatives should be readily evaluated using eq. (15), as all the other quantities in the equation, such as the vector, \mathbf{c} , and the Lagrange multiplier, λ , will have been determined in a preceding variational calculation.

CHOICE OF BASIS SET

The choice of forms for the basis functions, $\{f_i\}$, and for the charge density, ρ , and the dielectric constant, ε , will determine the efficiency and the precision of the basis set approach to the solution of the Poisson equation.

To start, a few general remarks on the choice of basis functions can be made:

- For the variational principle to be valid, the potential, and hence the basis functions, must tend to zero to infinity; that is $f_i(\mathbf{r}) \rightarrow 0$ as $\mathbf{r} \rightarrow \infty$.
- At large distances, the potential tends toward the same potential produced by the same charge distribution but in a dielectric constant equal to that outside the molecule. Therefore, the basis set should include functions that give this limiting behavior.
- The K_i integrals [eq. (9)] can be evaluated, regardless of the form of the basis functions, by transformation of the volume integral to that over a surface at infinity. For basis functions that have the limiting behavior of r^{-1} the integral will equal $-\varepsilon_0$ where ε_0 is the dielectric constant outside the molecule. For basis functions that decay more rapidly than this, the integrals are zero. Due to these results, some limitations are imposed on the choice of functions in a basis set. If the total charge on the system is nonzero then it is essential to use at least some basis functions with the r^{-1} limiting behavior; that is, with nonzero K_i integrals, so that the constraint

condition of eq. (8) can be satisfied. If the total system charge is zero then there is no such limitation.

In this initial work we have chosen to use Gaussian functions (or functions derived from them, such as error functions) as much as possible. This is because they are very amenable to various manipulations, including integrations, and because there exists a large body of experience from the quantum chemistry field on their use.⁹ Thus, instead of the point charge model that is usual for the charge density, we express the charge density of the molecule as a sum of Gaussian densities for each atom, ρ_α :

$$\rho(\mathbf{r}) = \sum_{\alpha=1}^N q_\alpha \left(\frac{1}{2\pi\sigma_\alpha^2} \right)^{\frac{3}{2}} \exp \left(-\frac{(\mathbf{r} - \mathbf{R}_\alpha)^2}{2\sigma_\alpha^2} \right) \quad (16)$$

where q_α is the atomic charge; σ_α is the width of the distribution for the atom, α ; and the number of atoms in the system is N .

For the basis functions we tried a number of forms which are all based upon the Gaussian and which are discussed more fully in the Results section. In particular, we used functions which represent the potential due to a Gaussian charge distribution; that is:

$$\int d\mathbf{r}' \frac{\rho_\alpha(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = q_\alpha \frac{\text{erf}(|\mathbf{r} - \mathbf{R}_\alpha|/\sqrt{2}\sigma_\alpha)}{|\mathbf{r} - \mathbf{R}_\alpha|} \quad (17)$$

$$= \sqrt{\frac{2}{\pi\sigma_\alpha^2}} F_0 \left[\frac{(\mathbf{r} - \mathbf{R}_\alpha)^2}{2\sigma_\alpha^2} \right] \quad (18)$$

where the function, $F_0[x]$, is given by:

$$F_0[x] = \int_0^1 \exp(-xs^2) ds \quad (19)$$

With these forms for the basis functions and the charge distribution the P_i integrals [eq. (5)] are readily evaluated.

The remaining problem is how to represent the dielectric function, $\varepsilon(\mathbf{r})$. It is well known that the volume of a molecule comprised of hard-sphere atoms can be obtained by integrating a "matter" density, $m(\mathbf{r})$, of the form^{14,15}:

$$m(\mathbf{r}) = 1 - \prod_{\alpha}^N (1 - \nu_\alpha(\mathbf{r})) \quad (20)$$

where ν_α is a step function centered at atom α which has the value of 1 inside the hard sphere

and a value of 0 outside. Thus, the matter density has a value of 1 inside the molecule and a value of 0 outside. We can use this function to define the dielectric function as:

$$\varepsilon(\mathbf{r}) = \varepsilon_i m(\mathbf{r}) + \varepsilon_0 (1 - m(\mathbf{r})) \quad (21)$$

where ε_i and ε_0 are the interior and exterior dielectric constants, respectively. To make the manipulation of this function easier we replace the step function form for ν_α by products of polynomials and Gaussians, the exact forms for which are discussed next. We note that Grant and Pickup have recently used single Gaussian functions in their calculations to determine the volumes and surface areas of molecules.¹⁵ With this form for the dielectric function, the F_{ij} integrals [eq. (6)], although not trivial, can, in principle, be evaluated relatively efficiently.

Results

SYSTEMS WITH SPHERICAL SYMMETRY

To obtain insights into the basis set approach it is useful to consider spherically symmetric systems with a single positive unit charge because the potential and the energy can be solved for numerically using standard techniques for arbitrary forms of the dielectric and charge distribution functions. The Poisson equation in this case reduces to a single variable equation:

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \varepsilon(r) \frac{d\phi}{dr} \right) + 4\pi\rho(r) = 0 \quad (22)$$

For the calculations in this section we represented the charge distribution as a single Gaussian function [eq. (16)] with $\sigma = 0.5$ Å. For the dielectric function we used the expression given in eq. (21) with values for the dielectric constants, ε_i and ε_0 , of 1 and 78.36, respectively. A number of forms for the function, ν , were tried:

$$\nu_{hs}(s) = \Theta(s - 1) \quad (23)$$

$$\nu_0(s) = \exp(-s^2) \quad (24)$$

$$\nu_2(s) = (1 + s^2)\exp(-s^2) \quad (25)$$

$$\nu_4(s) = \left(1 + s^2 + \frac{1}{2}s^4\right)\exp(-s^2) \quad (26)$$

$$\nu_6(s) = \left(1 + s^2 + \frac{1}{2}s^4 + \frac{1}{5}s^6\right)\exp(-s^2) \quad (27)$$

where Θ is the step function (with a value of 0 if the argument is greater than 0 and 1 otherwise).

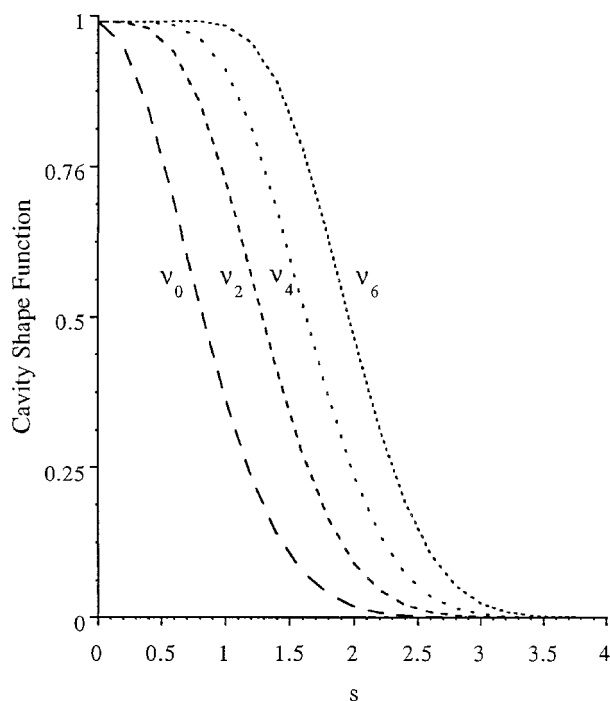


FIGURE 1. A plot of the functions used to represent the dielectric cavity for a single atom. The functions are defined in the text. The hard-sphere function (not shown) has a value of 1 up to $s = 1$ and 0 thereafter.

The parameter s is related to r by the expression $s = r/a$, where a is a parameter that determines the size of the cavity for the charge. These functions, except ν_{hs} , are plotted in Figure 1.

The *solvation energies* for the systems using the various forms for ν are shown in Figure 2 as a function of the parameter, a . The solvation energies are defined as the difference between the energy of the charge distribution in vacuum ($\epsilon_0 = 1$) and in the dielectric ($\epsilon_0 = 78.36$). It can be seen that the system with the dielectric cavity defined using the function ν_6 gives solvation energies that are in good agreement with the hard-sphere values. The energies for the systems using the other functions vary much less rapidly with distance. This is likely to be a problem in calculations on real molecules because very large a parameters would be necessary to obtain correct solvation energies. Therefore, for the rest of the calculations in this article, we chose to model the dielectric cavity for each atom using a ν_6 function.

To illustrate the basis set approach for the single atom system we present the results of a series of calculations using various basis functions in Table I. The a parameters were selected so as to mimic the solvation energies for the lithium and chloride

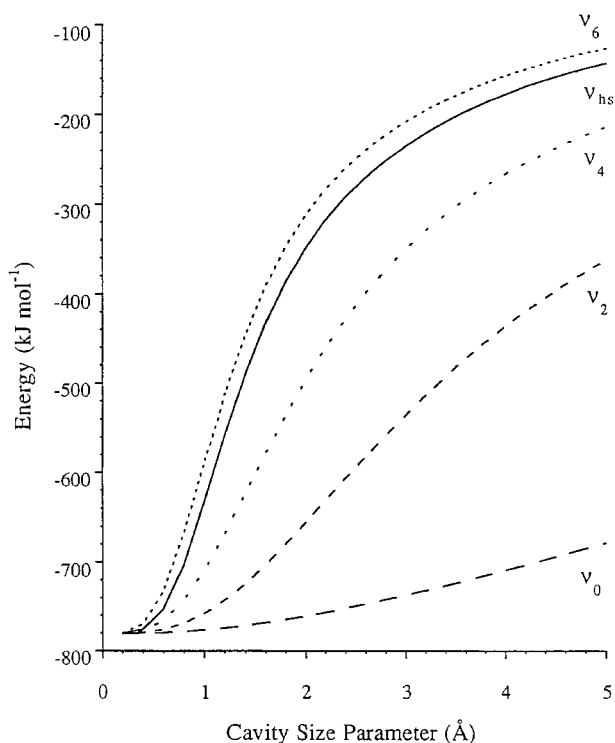


FIGURE 2. A plot of the solvation energy (kJ mol^{-1}) as a function of the cavity size parameter, a (\AA). The dielectric cavity functions are defined in the text and the interior and exterior dielectric constants were 1 and 78.36, respectively.

ions, respectively (i.e., $a = 1.2 \text{ \AA}$ and $a = 1.8 \text{ \AA}$). The basis functions are defined as follows:

$$f_l(r) = \frac{\text{erf}(r/\sqrt{2}\sigma)}{r} \quad (28)$$

$$f_{\nu_i} = \nu_i(r) \quad (29)$$

$$f_{\nu_i,l} = \nu_i(r) \times f_l(r) \quad (30)$$

where the $\nu_i(r)$ functions are defined in eqs. (24)–(27). All integrals using these functions were done numerically. Note that the value of the size parameter, a , used for the ν_i functions was the same as that used for the definition of the dielectric cavity (i.e., 1.2 \AA or 1.8 \AA depending on the calculation), whereas σ had the value 0.5 \AA .

It can be seen that it is possible to obtain good precision if a relatively small number of basis functions are used (three in this case), although the mix of functions is important. Thus, whereas the use of the functions, ν_2 , ν_4 , and ν_6 give poor results, even with three functions, ν_0 gives good results as long as three functions are used. It is noteworthy that the use of the functions,

TABLE I.
Energies Obtained with an External Dielectric
Constant of 78.36 by Solving the Poisson
Equation Using Various Forms for Basis
Functions for Single Particle System.^a

| Basis set | Energies (kJ mol ⁻¹) | |
|--------------------------|----------------------------------|-----------------------|
| | $a = 1.2 \text{ \AA}$ | $a = 1.8 \text{ \AA}$ |
| Exact | 277.0 | 442.7 |
| f_I | 3.8 | 4.1 |
| f_I, f_{I0} | 98.7 | 107.9 |
| $f_I, f_{I0, I}$ | 198.3 | 331.0 |
| $f_I, f_{I0}, f_{I0, I}$ | 263.0 | 427.6 |
| $f_I, f_{I2}, f_{I8, I}$ | 202.9 | 355.6 |
| $f_I, f_{I4}, f_{I4, I}$ | 135.1 | 249.8 |
| $f_I, f_{I6}, f_{I6, I}$ | 90.8 | 168.6 |

^aThe basis functions are defined in the text. The dielectric cavity function used was ν_6 with values for a of 1.2 and 1.8 Å, respectively. The value of the internal dielectric constant = 1.

$f_I(r)/\epsilon(r)$, as basis functions, give much worse results than the result cited in Table I, despite their additional complexity.

Instead of discussing further the symmetrical case we move on to the discussion of general molecular systems and see if an equivalent accuracy can be obtained.

MOLECULAR SYSTEMS

For the calculations in this section a set of six small molecules were selected. They were:

1. A diatomic consisting of two atoms 2 Å apart, each with a single positive charge and cavity size parameters of 1.5 Å.
2. The same diatomic but with one atom having a positive charge and the second a negative charge.
3. Methanol. The geometry and parameters for the molecule were taken from the examples in the test suite provided with the UHBD program.¹⁶ The OPLS charges and atomic radii were used¹⁷ and the atoms in the methyl group were treated as a single "united atom."
4. Formate anion. The parameters and geometry were obtained from the QUANTA molecular modeling program.¹⁸ The charges were 0, -0.628; C, 0.373; H, -0.117; the cavity sizes: O, 1.655 Å; C, 1.875 Å; H, 1.2 Å; and the geometry: C—O, 1.26 Å; H—C, 1.14 Å; H—C—O, 117.2°.

5. Acetic acid. The geometry and parameters for this molecule were taken from the UHBD program test suite, although, in this case, the methyl group was treated as four separate atoms.
6. Methazolamide. The geometry and parameters were taken from the UHBD example test suite.

The charge distributions of the atoms in all the molecules were represented by Gaussians with widths of 0.5 Å. Note that the variable parameters—the charges on the atoms, the width of the charge distribution, and the cavity sizes—were all used without optimization for these calculations. No attempt was made to vary the parameters to obtain realistic values of the solvation energy, for example. This will be the aim of future work if the basis set method appears to be competitive with the finite difference approaches.

The results of four sets of calculations, performed using different methods of solution, will be presented for these molecules. The methods are:

1. Finite differences (FD). To compare the results calculated using various basis sets we modified a finite difference program that we had written previously, which solves the Poisson equation.¹⁹ To do this, the values of the dielectric on the grid are assigned using the form of the dielectric function [eqs. (21) and (27)] and the charges for each grid point are calculated by integrating the charge density [eq. (16)] over a cube centered on each grid point. The method of iterative solution is the same, although, in passing, we note that convergence (as a function of grid size) was found to be much more rapid with a Gaussian distribution of charge than when using the more traditional representation of a set of point charges. The values of the self-energies found when using small grid sizes were also much reduced. For the small molecules we used a grid of size 100^3 with a spacing of 0.2 Å. For the larger molecules (formate, acetic acid, and methazolamide), we used a grid of size 120^3 with spacings of 0.18 Å, 0.186 Å, and 0.227 Å, respectively.
2. Basis set 1 (BS 1). This basis set consisted of $3N$ functions for each molecule (where N is the number of atoms in the molecule). Each atom was assigned three functions of the forms given in eqs. (28)–(30) with the ν_0 function [eq. (24)] being chosen for ν_i .

3. Basis set 2 (BS 2). This basis set consisted of *only* three functions for each molecule. These were:

$$f_T(\mathbf{r}) = \sum_{\alpha=1}^N \frac{q_{\alpha} \operatorname{erf}(|\mathbf{r} - \mathbf{R}_{\alpha}|/\sqrt{2} \sigma)}{|\mathbf{r} - \mathbf{R}_{\alpha}|} \quad (31)$$

$$f_{\nu}(\mathbf{r}) = 1 - \prod_{\alpha=1}^N (1 - \nu_{0_{\alpha}}(|\mathbf{r} - \mathbf{R}_{\alpha}|)) \quad (32)$$

$$f_{\nu,T}(\mathbf{r}) = f_{\nu}(\mathbf{r}) \times f_T(\mathbf{r}) \quad (33)$$

The choice of f_T ensures that the potential at long range has the correct form. This is because the constraint condition [eq. (8)] for a charged system can only be satisfied if the coefficient for this basis function has the value $1/\varepsilon_0$ (because all the K_i integrals are zero except for f_T).

4. Basis set 3 (BS 3). As a result of our experience with the previous two basis sets we combined the three functions of BS 2 with the $2N$ short-range basis functions [eqs. (29) and (30)] of basis set 1 giving $2N + 3$ basis functions for each molecule.

The electrostatic energies for the molecule in vacuum (i.e., with a dielectric of 1 everywhere) can, of course, be calculated analytically and provide a

check on the accuracy of the finite difference calculations. With the choice of the basis functions given above it is, in principle, possible to calculate the integrals P_i and F_{ij} analytically²⁰ (because all the functions are Gaussians). However, for this work we only evaluated the integrals analytically which did not involve the functions $m(\mathbf{r})$ or f_{ν} [eqs. (20) and (32), respectively]. For integrals involving these functions we used an adaptive numerical integration algorithm in three dimensions.²¹ We checked the values produced by the integration and believe they are accurate, in all cases, to 10^{-3} .

The results of the calculations are given in Table II. A number of points are worthy of note. First, in some cases, the energies for $\varepsilon_0 = 78.36$ obtained using the basis set methods are slightly larger than the finite difference results (they should, of course, be smaller). We have investigated this and are confident that it is due to numerical errors in the finite difference results. Increasing the number of grid points and decreasing the size of the grid improves the quality of the values, although those reported in Table II are at the limit of the computer resources that we had available. Second, it is possible to obtain results which are the same (to numerical accuracy) as those of the finite difference results using a reasonable number of basis functions (of the order of N). Third, and perhaps surprisingly, the results with a basis set consisting

TABLE II.
Energies Obtained for Some Molecular Systems by Solving the Poisson Equation Using a Finite Difference (FD) Method and a Basis Set Method with Three Different Sets of Basis Functions.^a

| Molecule | Diatomic | Diatomic | Methanol | Formate | Acetic acid | Methazolamide |
|-------------------------|----------|----------|----------|---------|-------------|---------------|
| <i>N</i> | 2 | 2 | 3 | 4 | 8 | 21 |
| Charge | 2 | 0 | 0 | -1 | 0 | -1 |
| Energies | | | | | | |
| $\varepsilon_0 = 1$ | 2259.1 | 876.3 | 128.4 | 555.9 | 246.1 | 1159.1 |
| FD | | | | | | |
| $\varepsilon_0 = 78.36$ | 992.7 | 658.1 | 110.0 | 313.0 | 229.5 | 916.7 |
| Solvation | -1266.4 | -218.2 | -18.4 | -242.9 | -16.6 | -242.4 |
| BS 1 | | | | | | |
| $\varepsilon_0 = 78.36$ | 940.9 | 657.4 | 90.6 | 303.4 | 230.2 | 913.2 |
| Solvation | -1318.2 | -218.9 | -37.8 | -252.5 | -15.8 | -245.9 |
| BS 2 | | | | | | |
| $\varepsilon_0 = 78.36$ | 989.6 | 637.1 | 110.6 | 312.6 | 229.9 | 886.8 |
| Solvation | -1269.5 | -239.2 | -17.9 | -243.3 | -16.1 | -272.3 |
| BS 3 | | | | | | |
| $\varepsilon_0 = 78.36$ | 992.6 | 657.7 | 110.7 | 315.0 | 231.3 | 918.5 |
| Solvation | -1266.5 | -218.6 | -17.7 | -240.9 | -14.8 | -240.6 |

^aThe basis functions are defined in the text. The internal and external dielectrics were 1 and 78.36, respectively, and all energies are expressed in kJ mol^{-1} . The solvation energy is simply the difference between the energies with exterior dielectrics of 1 and 78.36, respectively.

of only three basis functions per molecule (BS 2) are only significantly worse than those evaluated with the best basis set (BS 3) in the case of the neutral diatomic and methazolamide. This raises the possibility of being able to find a basis set method consisting of a small number of functions ($\ll N$) that can be used for calculations on large systems when the solution of the matrix equations for the coefficients [eq. (10)] with a large number of basis functions could become a limiting factor.

Conclusions

In this article we have explored the use of a basis set expansion approach for the solution of the Poisson equation for small molecules immersed in a solvent that is represented by a dielectric continuum. We have found that a basis set method, with a relatively small number of basis functions [$\sim O(N)$], can give results that are in very good agreement with those obtained by solving the Poisson equation using a finite difference method. To make the basis function method practical for application to larger systems we used functions of Gaussians to represent the charge distribution of the molecule and the cavity formed by the solute in the solvent.

The aim of the work presented here has been to demonstrate that the basis set expansion method is feasible for the solution of the Poisson equation. As a result we have not focused on issues such as the computer time taken by the algorithm or by the optimization of the program, although the calculations reported here, even with the numerical integration algorithm, took about the same amount of time as the finite difference methods. We expect these times to be significantly reduced when more advanced methods are employed.

The main objective of our future work will be to see if the method can be competitive with finite difference methods for applications to proteins and other macromolecular systems. In particular we will investigate alternative forms for the cavity function and the basis functions and look at ways in which the integrals involving these functions can be efficiently evaluated. We will also investi-

gate the parameterization of the basis sets. All these aspects will be discussed in another publication.

Acknowledgments

The authors thank the Institut de Biologie Structurale—Jean-Pierre Ebel, the Commissariat à l'Énergie Atomique, and the Centre National de la Recherche Scientifique for support of this work.

References

1. S. C. Harvey, *Prot. Struct. Function Genet.*, **5**, 78 (1989).
2. K. A. Sharp, *Curr. Opin. Struct. Biol.*, **4**, 234 (1994).
3. M. K. Gilson, *Curr. Opin. Struct. Biol.*, **5**, 216 (1995).
4. B. Honig and A. Nicholls, *Science*, **268**, 1144 (1995).
5. T. J. You and S. C. Harvey, *J. Comput. Chem.*, **14**, 484 (1993).
6. H.-X. Zhou, *Biophys. J.*, **65**, 955 (1993).
7. R. Bharadwaj, A. Windemuth, S. Sridharan, B. Honig, and A. Nicholls, *J. Comput. Chem.*, **16**, 898 (1995).
8. M. K. Gilson, J. A. McCammon, and J. D. Madura, *J. Comput. Chem.*, **16**, 1081 (1995).
9. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
10. K. A. Sharp and B. Honig, *J. Phys. Chem.*, **94**, 7684 (1990).
11. E. S. Reiner and C. J. Radke, *J. Chem. Soc. Faraday Trans.*, **86**, 3901 (1990).
12. M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon, *J. Phys. Chem.*, **97**, 3591 (1993).
13. T.-S. Lee, D. M. York, and W. Yang, *J. Chem. Phys.*, **102**, 7549 (1995).
14. K. D. Gibson and H. A. Scheraga, *Mol. Phys.*, **62**, 1247 (1987).
15. J. A. Grant and B. T. Pickup, *J. Phys. Chem.*, **99**, 3503 (1995).
16. M. E. Davis, J. D. Madura, B. A. Luty, and J. A. McCammon, *Comp. Phys. Commun.*, **62**, 187 (1991).
17. W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).
18. The QUANTA Molecular Modeling Program, Molecular Simulations Inc.
19. L. David, unpublished work.
20. Note that by "analytic" we mean that the integrals can be expressed in terms of well-known functions even if some of these, such as the error function or F_0 , must be calculated numerically.
21. The NAG Fortran Library, Mark 16, The Numerical Algorithms Group Ltd., Oxford, 1993.